

Tone-TBU Disparity in Three Phonologies of Cantonese

Luca Iacoponi
Rutgers University

Abstract

I propose a restricted theory of tone-TBU association called Basic Tonal Theory (BTT). BTT includes six well-attested constraints, and a restricted GEN. The typology contains attested cases of tone over-crowding resolution such as tone deletion, mora epenthesis, and mora deletion, but also predicts unattested cases of languages that resolve tonal crowding by using a combination of the aforementioned mechanisms based on the number of the morae in the input.

Cantonese has both monomoraic and bimoraic syllables, lexically specified tones, and a process of tonal suffixation that creates the context for tonal overcrowding. In an acoustic experiment, the duration and pitch of pairs of suffixed and underived forms were measured, and the analysis of the results shows that two of the three groups use a split-strategy to resolve tonal crowding.

The experiment thus provides empirical evidence for the typology of BTT. Furthermore, the proposed analysis can account for a case of near mergers previously argued to be unexplainable in an optimality framework.

1 Basic Tonal Theory (BTT)

1.1 Gen BTT is a restricted theory of tone-TBU association. It generates only one syllable at a time, in which (t) is used to indicate any tone and the input and output are as follows:

Input = $\{[t]\mu, [tt]\mu, [tt]\mu\mu, [ttt]\mu\mu\}$ Lexical and changed tones (CVO_{lex} , CVO_{tc} , CVR_{lex} , CVR_{tc})

Output $\in \{[t^+]\mu^+\}$ A sequences of 1 or more t followed by 1 or more u

Correspondence: $\langle in, out, corr \rangle$ with notation: α symbol, α' output, and \leftrightarrow ‘corresponds to’ which typed as $t \leftrightarrow t'$ and $\mu \leftrightarrow \mu'$. Each $x \in in$ corresponds with 1 and only 1 $x' \in out$, and each $x' \in out$ corresponds with 1 and only 1 $x \in in$.

1.2 CON The constraints (CON_{BTT}) are:

f: max-t $t \in in \Rightarrow \exists t' \in out \text{ s.t. } t \leftrightarrow t'$

f: dep-t $t' \in out \Rightarrow \exists t \in in \text{ s.t. } t \leftrightarrow t'$

f: dep- μ $\mu' \in out \Rightarrow \exists t \in in \text{ s.t. } \mu \leftrightarrow \mu'$

M: $\sigma = \mu\mu$ $card \{\mu \in out\}$
 σ contains exactly two μ

M: noCompC $card \{\mu \in out\} - 2\sigma$
contains at most two t (Yip, 2002)

M: noShortT $card \{t \in out\} - card \{\mu \in out\}$
for each μ in the output there is at most one t in the output (Yip, 2002)

* I would like to thank Roxana Fung, Ken Cheng, Ou Jinghua and Kwok Ciu Lai for their help and support at PolyU, 赵子 at Jinan University in Guangdong for offering me access to the laboratory, finding me the speakers and for being an exquisite host, and indeed all who volunteered to take part in the experiment. I would also like to thank Paul de Lacy for enormously improving this handout with structural and substantial suggestions, Alan Prince for pointing out theoretical issues and making this handout clearer, Aaron Braver for sharing with me his expertise on incomplete neutralization. I am also grateful to Giovanna Marotta, Alan Yu, William Bennet, Eric Baković and all the attendees and organizers of Phonology 2013 for the comments and suggestions.

Other worthy-mentioning constraints that determine TBU-Tone association are **M:specifyT**, **M:nolongT**, and **F:NoContour.M:specify** requires at least one tone to be specified per syllable. This constraint is embedded in GEN in the current model. **M:nolongT** (C.F. **M:noshortT**) is irrelevant for tonal overcrowding: card $\{\mu \in \text{out}\} - \text{card } \{t \in \text{out}\}$.

In the theory, the morae are part of the input. However, the system is also an abstraction of models where the moraic structure is introduced in the output. The absence of a constraint on mora deletion (**f:max- μ**) makes the system a valid approximation, so long as segmental changes are not considered.

The typology generated by this set of constraints is a sub-set of BTT's. BTT predicts that tonal crowding can be resolved via moraic epenthesis, deletion of the root tone, faithfulness, tone shortening, or downstep (fig. 1). Those cases are well attested. The theory, though, also predicts peculiar cases where a different strategy can be chosen in a language depending on the different number of tones and TBU in the output (split strategy languages). For example, a language may choose to epenthesize a mora in monoraic domains with one extra-tone, and then leave the same tone unparsed in case of bimoraic input with two tones. These cases are hard to find evidence for, since the language has to show tonal crowding, as well as showing 1-to-1 tone-TBU association in monomoraic and bimoraic syllables. In the next sections I will show that Cantonese satisfies these requirements, and that two varieties of Cantonese are actual instantiation of split strategy languages.

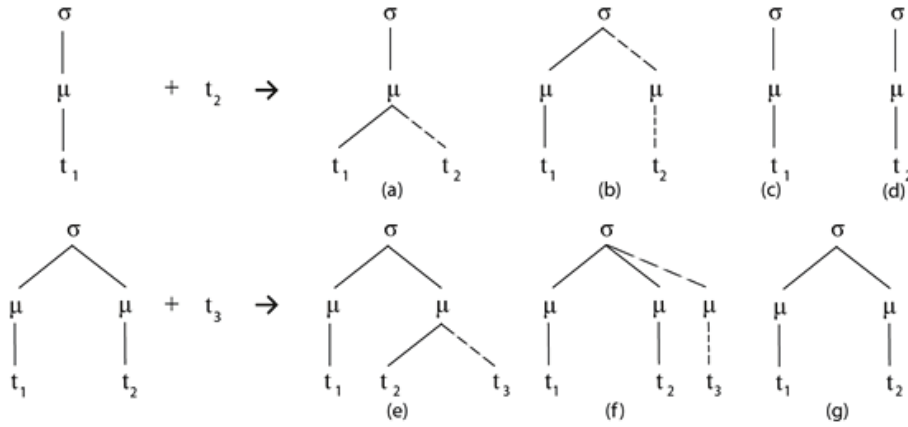


Figure 1: Possible resolutions of tonal crowding

2 Cantonese

2.1 Tonomatics Cantonese is a Sinitic language of the Yue family which is spoken in the East Guangdong/West Guangxi region, in Southern China, Hong Kong, Macau, and in most Chinatowns worldwide. It is the native language of approximately 60 million people² (Lewis et al., 2013). It is widely documented and attested in both theoretical and descriptive literature (see Bauer and Benedict, 1997 for a review).

There are nine tones of contours consist of 3 levels, 3 contours and 3 long level tones. The 3 levels are L, M, and H. The 3 contours are LM, MH, and ML. Three long level tones: HH, MM, and LL. Cantonese distinguishes between two syllable types CVO and CVR (Chao, 1974; Bauer and Benedict, 1997 and other). CVO syllables end in unreleased stops (/p, t, k/) and they are shorter than long syllables. Contours occur only in TC tones (HM).

CVR syllables end in sonorants. They are longer than CVO syllables and lexically specified with a contour or a long level tone (HH, MM ... LM, MH). The TBU is the mora in Cantonese (Iacoponi 2013). Codas are moraic in CVR syllables, but non-moraic in CVO syllables. Evidence from processes that changes only one tone in the syllable (see for instance 2.3) shows that long level tones are composed of two level tones. The tonal inventory as described in table 1. The structure proposed in (7-10).

² This number actually refers to all Yue dialects. The speakers of Cantonese *stricto sensu* are about 10 million.

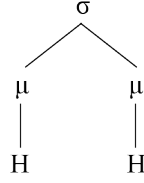


Figure 2: Representation of an HH tone in Cantonese

IPA+Chao	Gloss	Character	Tone Letter	
sik ⁵	“style”	式	H	CVO
sik ³	“to stab”	刺	M	
sik ²	“to eat”	食	L	
si: ⁵⁵	“to think”	思	HH	CVR
si: ³⁵	“history”	史	MH	
si: ³³	“to try”	試	MM	
si: ²³	“city”	市	LM	
si: ²²	“fact”	事	LL	
si: ²¹	“time”	時	ML	

Table 1: Tones of Cantonese

2.2 High Rising Tone Change The term high raising tone change (TC) is used to describe an array of processes that due to the grammar of Cantonese all result with the same output (Iacononi, 2013). It is one of the few tonal alternations in the language. In high raising tone change (TC) all tones change to a tone similar or identical to lexical MH (MH_{tc} vs MH_{lex}), with an exception of high tones, which are claimed to remain unchanged (Bauer, 1997). The array of processes that forms TC is:

Red/Prefix:

(1) /gwaɪ^{MM} **gwaɪ^{MM}H** + dei^{MH}/ → [gwaɪ^{MM} **gwaɪ^{MH}** dei^{MH}] “It’s a bit expensive” 贵贵咗

c.f. [gwaɪ^{MM}] “expensive”

Suffix (voc):

(2) /a^{MM} **tsiu^{LL} + H**/ → [a^{MM} **tsiu^{MH}**] “Chiu (voc/suffix)” 阿赵

c.f. [tsiu^{MM} sin^{HH} sang^{HH}] “Mister Chiu” 赵先生

Suffix (diminutive):

(3) /dou^{MH} **se^{MM} + H**/ “to pour out” → [dou^{MH} **se^{MH}**] “to spill” 倒瀉

Syllable Elision:

(4) /ŋɔ^{LM} **pək^M** tɕɔ^{MH} li^{MM} ɔɔ^{MM}/ → [ŋɔ^{LM} **pək^{MH}** li^{MM} ɔɔ^{MM}] “I chopped this (elision)”

c.f. [ŋɔ^{LM} pək^M tɕɔ^{MH} li^{MM} ɔɔ^{MM}] “I chopped this (slow speech)”³

³ Only H tones can trigger TC, see Iacononi (2013) for a discussion.

Notice that in (4) a CVO (CVO_{tc}) syllable surfaces with a contour. CVO_{lex} is distinguished from CVO_{tc} as the latter surface as a contour tone. There are two aspects of tone change; the first is change in the structure of the tone root (by adding or removing tones or TBU). The second aspect is change in the tone values of the input syllable (tonomatics restrictions, tone sandhi etc.). Since BTT does not distinguish between tone levels, and does affect tone-TBU association, the latter will not be discussed (see Iacononi, 2013a for an analysis).

3 Methodology

3.1 Questionnaire A first list of more than 500 actual words was compiled from diversified sources. The list was refined by two informed speakers of the Guangdong and the Hong Kong Cantonese. Non frequent forms, archaisms, and cases where TC was never possible were taken out the list.

The following criteria were used to reduce the list size and maintain phonetic homogeneity on the rhyme (except for elision context): /a/ plus nasal or /i/ or /u/ coda (CVR such as an, ai, au). Words rhyming in ət, ok, ək (CVO syllables). HH forms in Table 2 were used to investigate an unrelated question.

	CVR		CVO	total
tone	MH	HH		
TC	54	14	47	115
control	10	10	10	30
total	64	24	57	145

Table 2: syllable types used in the experiment

3.2 Sampling process. The conditions for participating in the experiment were the following: The participant must be born and raised in Guangzhou city (gz.), Hong Kong city (hk.). He/she must have native fluency in his/her variety. He/she must hold a high school diploma as minimum education requirement. The participant mustn't have any hearing, visual or cognitive impairments. All speakers should be fluent in at least two languages (English and Cantonese, Mandarin and Cantonese or all the three languages). Lastly, at least one of the parents must satisfy all the criteria above.

ID	Name	Age	Sex	City	District
1	WI	70	M	Hong Kong	Kowloon
2	PA	40	M	Hong Kong	Kowloon
3	LA	25	F	Hong Kong	New Territories
4	NG	45	M	Hong Kong	Kowloon
5	ZH	25	M	Guangzhou	Downtown

Table 3: gender, age and location of the participants⁴

⁴ It might be relevant at this point to point out that any sociolinguistic conclusion based on this data would be unfounded. Contrary to the common practice, the speakers were chosen as to reflect the highest possible variability. The interest was in finding the possibility of individual speakers' grammars, and not in pursuing any generalization about how those grammars are partitioned among the population.

3.3 Material and Procedure A Shure SM81-LC cardioid condenser microphone was used with Frequency range of 20-20,000 Hz, nearly flat response and Self-noise of 16 dB typical. A laptop equipped with a Conextant 20671 sound card supporting 24 bit resolution, a 24 bit Audiobox22VSL pre-amplifier, and 0.5 meters cable with 100% shielding factor were also used. Sample rate was 44100Hz.

For the recordings, the distance between the speaker and the microphone was 10cm. The participants were tested one at a time, and they were given the opportunity to decline to participate in the study. Between each reading page, participants were free to take a short pause of some seconds. All recording were made in the silent room of the Laboratory of Phonetics in Guangzhou at Jinan University 暨南大学, and calm and silent environment of High quality Audio-Visual Laboratory at the Hong Kong Polytechnic University.

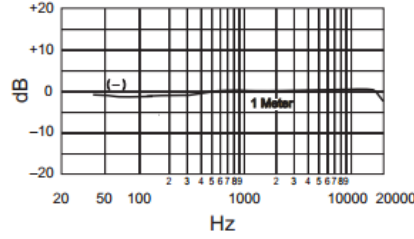


Figure 3: Frequency response of the microphone used for the recording (adapted from the SM81-LC spec sheet)

Pitch contours were extracted using the option pitch listing in Praat v. 5.3.18. Pitch range settings were 75-300 Hz for man and 100-500 Hz for women. In order to normalize contours, the following formula was used (Zhu, 2010):

$$z_i = \frac{\log_{10} x_i - \frac{1}{n} \sum_{i=1}^n \log_{10} x_i}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n \left(\log_{10} x_i - \frac{1}{n} \sum_{i=1}^n \log_{10} x_i \right)^2}}$$

The highest peak in the first part of the rhyme (*tone onset*), the *fall* and the highest pitch in the final part of the syllable (*peak*) were measured (figure 4). In fast speech, or with phonetically continuous segments, it is impossible to determine *tone onset*, so it was not measured in these cases (figure 5). The measurements of tone onset are only corollary to the main point of the analysis, which will be based on peak f_0 and duration.

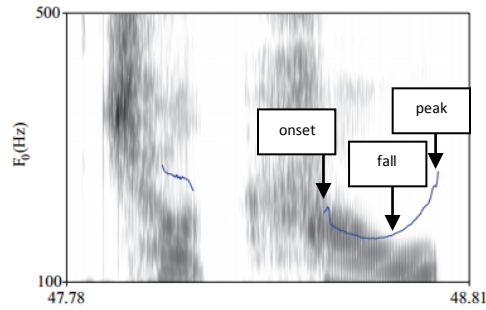


Figure 4: tone onset, fall and peak measure points

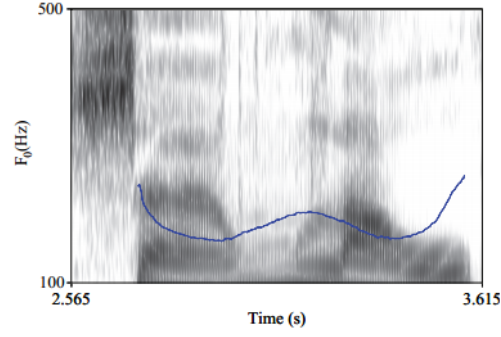


Figure 5: *tone onset* is impossible to determine in the context of two continuous segments

4 Results

4.1 Experimental Results The experimental results concerning CVR syllables were as follow: MH_{lex} was found to be different from MH_{tc} (MH^+) in both f_0 and duration of the rhyme for 3 out of 5 speakers (WI, PA, LA). MH_{lex} is identical to MH_{tc} for the other two (ZH, NG). The changed tone for the speakers that distinguish it starts lower and ends higher than the lexical tone (figure 6). Also, this data confirms a similar experiment (Yu, 2007). The experimental results concerning CVO syllables were as follow: For WI there is no difference in duration for CVO_{tc} and CVR syllables. For all other speakers, there is a significant difference between the three tones, as represented by the scale: $CVO_{lex} < CVO_{tc} < MH_{lex}$. The speakers can be divided into the following three groups:

	MH^+	$CVO_{tc} = CVR_{lex}$
G1	×	×
G2	✓	✓
G3	✓	×

Table 4: super high pitch and syllable duration for the three groups

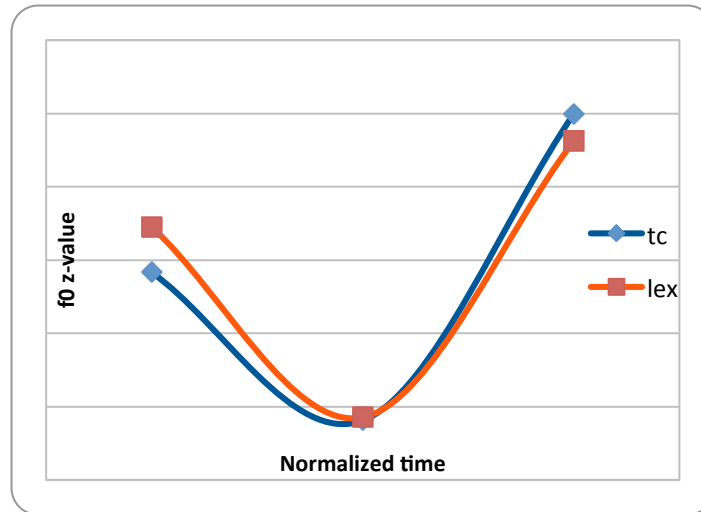


Figure 6: MH_{lex} and MH^+ average normalized pitch curve

4.2 Analysis There are three instances of the typology in which we can abstract the aforementioned scale to indicate a difference in either pitch, duration or both and link it to their structural description and constraint ranking⁵. G1, G2, G3 correspond to three instances of the typology.

G1 is the first example of split strategy. In CVO, the suffixal H tone adjoins to mora, resulting in a CVO_{tc} longer than CVO_{lex}, but crucially not as long as MH_{lex}. If the mora was epenthesis, as in G2, we would in fact expect CVO_{tc} and MH_{lex} to be almost identical (module the shorter duration of the obstruent). The speakers of G1 instead delete the second root tone to realize the suffixal tone, and therefore do not distinguish MH_{lex} from MH_{tc}, as they both have the same contour and the same number of morae. The phonetic hierarchy aligned with the phonological representations is showed in (5). Grammar ranking and support are in (6). (7) includes the acoustic measurements.

$$(5) \quad \begin{array}{ccccc} \text{CVO}_{\text{lex}} < & \text{CVO}_{\text{tc}} < & \text{MH}_{\text{lex}} = & \text{MH}_{\text{tc}} \\ \text{M}\mu < & \text{MH}\mu < & \text{MH}\mu\mu \end{array}$$

$$(6) \quad \mathbf{f: max-\mu, f: dep-\mu \gg f: max-t, M: \sigma=\mu\mu \gg M: noCompC, M: noShortT}$$

Input	Winner	Loser	f:Max-μ	f:Max-t	M:σ=μμ	f:Dep-μ	M:noCompC	M:noShortT
[tt]μ	[tt]μ	[tt]μμ		W			L	L
[ttt]μμ	[tt]μμ	[ttt]μμ			W	L		W
[tt]μ	[tt]μ	[t]μ				W		L

(7)

	duration (ms)			pitch (z-value)				
	CVO _{lex}	CVO _{tc}	p-value	MH _{lex}	p-value	MH _{lex}	MH _{tc}	p-value
ZH	114	174	<0.01	229	<0.01	0.98	1.07	0.78
NG	94	217	<0.01	237	<0.01	1.05	1.14	0.52

G2 is the second example of split strategy observed, with mora epenthesis in CVO_{tc}, and tone adjoined to the second mora in CVR_{tc}. CVO_{tc} is not significantly different from MH_{lex}, thus suggesting that they have the same phonological structure (9) and therefore that the second mora was epenthesis. G2 speakers shows the first case what has been analyzed as a case of incomplete neutralization (Yu, 2007). MH_{lex} is in fact slightly shorter than MH_{tc}. The phonetic difference reflects a structural difference, and it is therefore a possible language in OT, as it is in fact generated by BTT (10) (contra Yu, 2007).

$$(8) \quad \begin{array}{ccccc} \text{CVO}_{\text{lex}} < & \text{CVO}_{\text{tc}} & = & \text{MH}_{\text{lex}} < & \text{MH}_{\text{tc}} \\ \text{M}\mu < & \text{MH}\mu\mu & & < & \text{MMH}\mu\mu \end{array}$$

⁵ How pitch and duration correlate is unclear and problematic, and so it will be not discussed here.

(9) **f: max-μ, f: max-t, M: σ=μμ > f: dep-μ, M: noShortT, M: noCompC**

Input	Winner	Loser	f:Max-μ	f:Max-t	M:σ=μμ	f:Dep-μ	M:noCompC	M:noShortT
[ttt]μμ	[ttt]μμ	[t]μμ		W			L	L
[tt]μ	[tt]μμ	[tt]μ			W	L		W
[ttt]μμ	[ttt]μμ	[ttt]μμμ			W	W		L

G3 is not a split-strategy grammar. The suffixal tone always adjoins to the root mora, no matter the number of mora-bearing segments in the input. This results in all output forms to be phonetically and phonologically different (11), as adjoining to mora will result in different structure for CVO_{tc} and MH_{lex}. Ranking grammar and support are given in (12).

(10) CVO_{lex} < CVO_{tc} < MH_{lex} < MH_{tc}
Mμ < MHμ < MHμμ < MMHμμ

(11) **f: max-μ, f: dep-μ > f: max-t, M: σ=μμ > M: noCompC, M: noShortT**

Input	Winner	Loser	f:Max-μ	f:Max-t	M:σ=μμ	f:Dep-μ	M:noCompC	M:noShortT
[ttt]μμ	[ttt]μμ	[ttt]μ	W			W		L
[ttt]μμ	[ttt]μμ	[t]μμ		W			L	L
[tt]μ	[tt]μ	[tt]μμ			W	L		L

(12)

	duration (ms)				duration (ms)			
	CVO _{lex} (ms)	CVO _{tc}	p-value	MH _{lex}	p-value	MH _{lex}	MH _{tc}	p-value
PA	84	166	<0.01	231	<0.01	0.93	1.14	<0.01
LA	161	278	<0.01	314	<0.01	0.81	1.0	<0.01

4.3 Summary Three different groups of speakers are found with respect to how contrast is realized on lexical vs TC syllables. These three groups differ in the way the same input is mapped into different surface structures. G3 groups with G1 for CVO syllables and with G2 for CVR. G1 and G2 are two cases of split strategies, where monomoraic and bimoraic syllables different in the strategy adopted to resolve tonal crowding. G2 and G3 are both instances of grammars previously analyzed in Cantonese as cases incomplete neutralization. (11) shows the different phonological structure for G1, G2 and G3 in the four contexts analyzed (CVO_{lex}, CVO_{tc}, CVR_{lex} and CVR_{lex}), as well as providing the hasse diagrams for each ranking grammar.

(13)

	CVO _{lex}	CVO _{tc}	CVR _{lex}	CVR _{tc}	c-ranking
G1	σ μ t	σ μ / \ M H	σ / \ μ μ M H		
G2	σ μ t	σ / \ μ μ M H		σ / \ μ μ L/M M H	
G3	σ μ t	σ μ / \ M H	σ / \ μ μ M H	σ / \ μ μ L/M M H	

5 Conclusions

I proposed a theory of tone-TBU association called Basic Tonal Theory (BTT). BTT generates a typology that includes languages where tonal crowding is resolved differently in the same grammar depending on the number of mora-bearing segments in the input. By coupling experimental and phonological analysis, it was showed that two instances of such languages are found in Cantonese. Furthermore, all the phonetic differences between the three groups of speakers have been demonstrated to result from differences in phonological representation, contrary to the analyses of incomplete neutralization.

References

- Bauer, Robert S. & Benedict, Paul K. (1997), *Modern Cantonese Phonology*, Berlin: Mouton de Gruyter.
- Braver, Aaron & Kawahara Shigeto. (2013), 'Complete and Incomplete Neutralization in Japanese Monomoraic Lengthening', ms. Rutgers.
- Chao, Yuen Ren. (1956), 'Formal and semantics discrepancies between different levels of Chinese structure', *Bulletin of the Institute of History and Philology* **28.1**, 1–16.
- Chen, M. Y. (2004), *Tone Sandhi: patterns across Chinese dialects*, 'Cambridge studies in linguistic', Cambridge University Press.
- Iacoponi, Luca. (2013a), 'High Tone Mergers and Tone Change: Positive Conspiracy in Cantonese'. [Poster]. Annual Conference of the International Association of Chinese Linguistics (IACL). Taipei: Taiwan.
- Iacoponi, Luca. & Cheng, S. K. (2013a), 'Historical Affix Lenition and OT: The Case of Yue Dialects' Diminutive Suffix. [Handout]. Annual Conference of the International Association of Chinese Linguistics (IACL). Taipei: Taiwan.
- Iacoponi, Luca. (in progress), 'A domain-theory of tonal scarcity and tonal crowding in Optimality Theory'.

-
- Prince, Alan. & Smolensky, Paul. (1993:2004), 'Optimality Theory: Constraint interaction in Generative Grammar'.
- Whitaker, K. P. K. (1955), 'A study on the modified tone in spoken Cantonese', *Asia Major* **5**, 9–36.
- Yip, Moira. (2002), *Tone*, 'Cambridge textbooks in linguistics', Cambridge University Press,
- Yu, Alan C. L. (2007), 'Understanding near mergers: the case of morphological tone in Cantonese', *Phonology* **24:1**, 187–214.
- Yu, Alan C. L. (2010), Tonal Mapping in Cantonese Vocative Reduplication, in 'Proceedings of the 35th Annual Meeting of the Berkeley Linguistics Society'.
- Yu, Alan C. L. (2011). Contrast Reduction. In John Goldsmith, Jason Riggle, and Alan C. L. Yu (eds.) *The Handbook of Phonological Theory*, Second Edition, chapter 9, pp. 291–318. Blackwell.
- Zhu, Xiaonong. (2010), 語音學 '*Phonetics*', Beijing: The Commercial Press.